Ontological Classification for Heritage Computer Aided Design

Noha Saleeb
Rene Garcia
School of Science & Technology
Middlesex University
London, UK
n.saleeb@mdx.ac.uk

Mohamed Marzouk
Faculty of Engineering
Cairo University
Cairo, Egypt
mm_marzouk@yahoo.com

***Abstract*—Recent endeavors in Computer Aided Design (CAD) have enabled creating intelligent 3D models of building assets that allow high efficiency in constructing and operating/maintaining those assets. However, in order to support this, detailed specification, classification and codification of the individual building systems and their components must be created. This entails usage and embedding of specific construction classification systems with the CAD systems utilized. While many such systems exist for new builds, there is lack of classification systems for heritage buildings, which are unique in the characteristics of their components that cannot be covered by current classification systems. This paper presents part of an ongoing research to justify and create a new Ontological Classification system for Heritage assets that can be used in conjunction with CAD systems, specifically 3D intelligent authoring tools, to specify the special requirements of Heritage components. The paper will start with a comparative study to validate creation of an "Onto-Classification" system as opposed to other systems like thesauri, taxonomies etc., including case study examples of them. This will be followed with examples of the new Onto-Classification system using a current existing Case study of Toson Palace in Egypt.**

Keywords—Compter Aided Design, Classification, Heritage BIM, Ontology, Taxonomy, Thesaurus

I. INTRODUCTION

It is crucial within the Engineering and Construction industry to utilize classification systems to specify and codify the different components and systems within a building for purposes of ordering, constructing and maintenance. This becomes even more relevant when considering historical or heritage buildings, which belong to different architectural styles and historical eras, built in different unique methods and contain different building elements and components. The need to renovate, refurbish, and maintain them dictates accurate recording of their constituent building elements and blocks and their detailed characteristics to be able to replace or conserve them in the optimum methods possible.

Current classification systems utilized for new buildings and assets include CI/SFB, CAW, SFG20, Omniclass, Uniclass, NRM versions etc. However, Saleeb et al. [1] provide evidence for the lack of appropriate current classification systems for Heritage which lack dimensions required for classifying heritage components in terms of object types, hierarchy of tables/schedules, depth levels and appropriateness for different architectural styles and parametric geometries (e.g. origin, material, allowed stresses, proportions).

Saleeb et al. [1] concluded that the factors and dimensions lacking from current classification systems necessitate development of a new system for Heritage. Four main requirements for development were identified, in addition to further attributes essential for defining heritage components, which include both geometric and non-geometric information e.g. architectural style, geometric characteristics and ratio, condition, construction method, cultural value, material, color, reflectance characteristics.

Furthermore, the type of grouping proposed to classify the components was the Combinatory (faceted) grouping where classes of objects can be identified using multiple sets of attributes. A facet acts as a set of similar properties such as functions to enable categorizing objects accordingly. In a faceted classification, new objects can be continuously added [2]. This can be more suited for a heritage classification system as new found and unique objects may need to entered into the system constantly and which may be categorized using multiple attributes related to function, social value, environmental context etc., which are not conventional attributes in current classification systems. This is different from the current Direct (hierarchical) grouping where classes of objects are identified through a combination of properties; however, new objects cannot be accommodated without creating new revisions of the classification [3].

Due to this discrepancy in system requirements above, the next stage is to compare and contrast the current classification systems including thesauri, taxonomies and ontologies to determine the most appropriate system to use for Heritage assets, as detailed in the next section.

II. CLASSIFICATION VS. THESAURUS VS. TAXONOMY VS. ONTOLOGY

A. Analyzing Classifications and Thesauri

Is creating a Classification system sufficient for defining Heritage asset components? According to Miller [4], both Classification and Thesaurus schemes are tools used for indexing and retrieval of information, however there are a few differences between them.

1. Classification deals with organizing information mono-hierarchically according and limited to a single aspect or factor at a time. Every concept is dismembered and included in some categories. However, a thesaurus in principle is a poly-hierarchical system offering access to information via multiple interrelated aspects – “a vocabulary of a controlled indexing language” [5]. This means that a term or its synonyms can appear in more than area showing interconnections between different words. This is not a required function when classifying building objects where every component just needs to be uniquely identified and specified in an organized hierarchy of categories without showing how it is similar to other terms or connected to other objects.

2. Thesauri record a set of terms (words or phrases) covering some knowledge domain, with three types of relationship - equivalence, hierarchical and associative - between them [6]. Classification systems do not necessarily exhibit the equivalence (synonymy) between different terms or inter relations and associations between child components (mainly hierarchical parent relationships). While displaying inter-relationships is a useful functionality, it is not required when codifying and specifying different building components individually for procurement and maintenance purposes. E.g. relationship between a volute and corniche is important to know in certain contexts that study architectural and structural relationships, but is not necessarily conducive to specifying the different objects individually for monitoring or procuring.

3. A thesaurus is a classification based on natural-language words rather than abstract categories, it does not form a strict tree structure, and one term may have several “parents” at the level above [6]. Thesauri are fundamentally linguistic, while classification schemes organize conceptual categories. Thesauri find compact words or phrases to describe objects. With classification schemes, the goal is to have completely distinct conceptual categories that are mutually exclusive and jointly exhaustive. Classifications are generally further organized in a structured manner than thesauri [7]. For classifying building components, semantics and meanings of the terms is not the main focus. However, a strict tree hierarchy of components is crucial e.g. space→system→item.

4. Categorial analysis is based on categories constructed beforehand but clusters are created during an analytical process. Terms can be simultaneously included in several categories but in one cluster only. Therefore, categorial analysis can be deemed as a thesaural method and cluster analysis a classification method [8]. Clusters of terms should be mutually exclusive, i.e., no term in one cluster should appear in any other clusters without plausible cause [9]. Having several parent classes for one item is not useful as it would render codifying a particular component difficult. Mutual exclusivity of objects is building assets is important to avoid confusion and non-precision in procuring objects. As an example, a “corniche” might be part of “columns” category or “decorations” or “non-structural elements” category. However, when codifying a corniche, it might be preferable to have it as part of one class only for replacement or refurbishment techniques purposes.

5. While the same thesaurus term could be linked to more than one class number, a preferred place is selected for the concept in the schedules and a cross-reference made, in the form of hierarchical or associative relationships between the preferred and non-preferred location. The expression of a “preferred place” is a classification-based way of thinking. The relationships between preferred and non-preferred terms are not hierarchical or associative but only that of equivalence [10]. The link between lead-in (cross reference) and preferred term must be treated as a many-to-many relationship. Rows or records in tables have one or more key fields which guarantee their uniqueness, and links between records in different tables are represented by matching key fields. If this system is used for classifying objects, this requires normalizing the database so that each entity and relationship is stored only once, thus avoiding the problems of redundancy and possible inconsistency [6].

B. *Analyzing Classifications and Taxonomies*

Classification is "systematic arrangement in groups or categories according to established criteria." The term is comprehensive that incorporates any type of grouping according to criteria. However, a Taxonomy is the process of giving names to objects or groups of objects according to their positions in a hierarchy e.g. orderly classification of plants and animals according to their apparent natural relationships. The items are defined according to their relationship with the other items in the hierarchy [11]. With taxonomies, the hierarchical relationships usually rely more on internal characteristics inherent within the items themselves e.g. species, however with classifications, criteria can be defined based on any external factors, which is more relevant to classifying or organizing building components based on many different external factors e.g. building discipline, energy usage, structure, or in case of heritage, architectural era, function, dimensions etc.

Taxonomies are also more concerned with providing exhaustive lists while classification is not exhaustive. This is useful to be able to add new building components to the classification.

Taxonomies are based on providing a hierarchical relationship map between a multitude of items while classification usually only groups items according to one or two attributes. The fundamental difference is that taxonomies describe relationships between items while classification simply groups the items [11]. This is beneficial for defining a clear specification and codification of asset components.

C. *Analyzing Taxonomies and Ontologies*

An Ontology is concerned with highlighting the metadata of associative relationships between objects. It specializes in relationships and the intricacies between them. Taxonomy identifies relationships between items and categories, but lacks the complexity that ontology provides in terms of displaying the metadata of those items that can ultimately change the associations between them. Ontology is a collection of numerous taxonomies that can be used to describe a domain of knowledge along with the relationships among them [12].

As can be seen in figure 1, an ontology delves into describing the inter-relations between the different items of the networked hierarchy of elements, and not just define its presence. Applying metadata to the relationship itself is a very beneficial aspect provided by Ontology. This can be especially useful in Heritage even more so than creating new builds due to the historical, cultural, human and environmental contexts of the heritage asset. For example, connecting a type of window to a façade can have differing relationships based on the architectural period, location and cultural aspects (e.g. privacy considerations of that era). This means that a relationship could be conditional, temporary or seasonal.

This is inference, and is one of main features of ontology. Other Metadata for that relationship, such as date range, origin, material, allowed stresses, proportions, architectural style, geometric characteristics and ratio, condition, construction method, cultural value, colour, reflectance characteristics can be added.

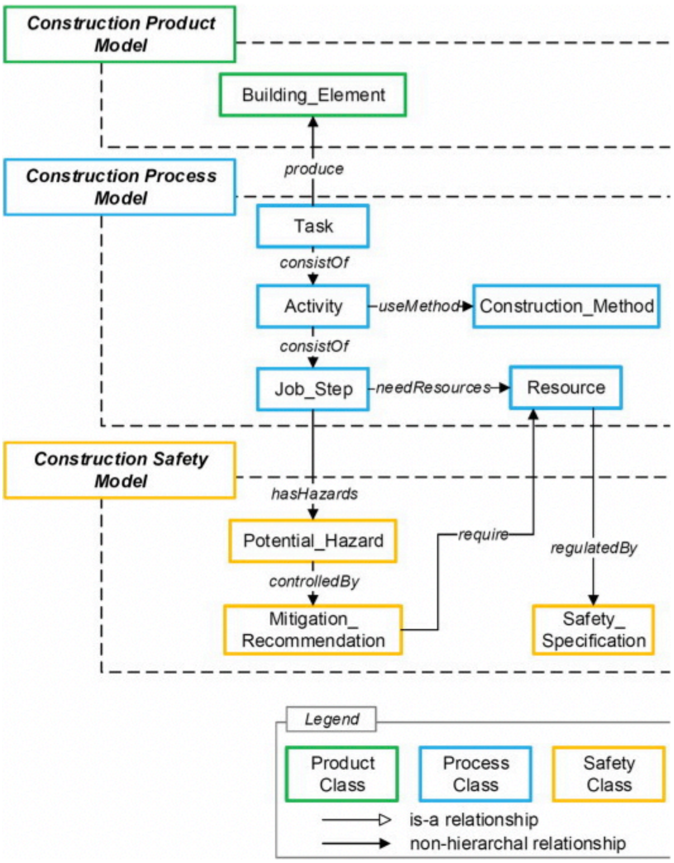

Fig. 1. Example of a construction related ontology [12]

Thus, the relationships and associations are not absolute. The different ways a relationship can be described and how that relationship may have facets, like seasonality, demonstrates the intricacy of an ontology. It makes the relationship active or inactive, which then triggers other relationships. While a taxonomy is a defined, static entity, an ontology is dynamic [13]. This could be perceived as a fundamental difference between live history and context of a Heritage building that affects how its components are refurbished and maintained, as opposed to a new build.

In the realm of Product Information Management (PIM), which can be similar to classification of building components or products, objects are either linked or not. Ontologies can provide added layers to that relationship and take it outside of PIM. Ontology looks at a much larger universe. "There can be all kinds of taxonomies in an ontology, says Dino Eliopulos, Managing Director at EIS, but the real difference is that an ontology attempts to describe and capture an entire subject area, with all of its complexity, whereas a taxonomy tries to simplify a complex collection of seemingly unrelated items into a linear, organization." [13]

III. DERIVING A NEW ONTOLOGICAL CLASSIFICATION SYSTEM

Based on the analysis of the different classification schemes described previously, the authors propose the adoption of a merged system to classify Heritage buildings within the context of BIM, comprising of a Classification system in principle which is enriched by adding aspects of associated metadata of the different components to it, as utilized within Ontologies. Hence the term an "Onto-Classification" is coined within this research i.e. a merged classification and ontology scheme. A summarization for the reasons for this is that a scheme is required to be

1. Non-exhaustive - allowing addition of new elements to it as previously explained in the introduction chapter. This is an attribute of classification systems as opposed to taxonomies.
2. Non-semantic specific – focus is not on meaning of words and which terms can be synonymous with each other, which is an attribute of Thesauri
3. Doesn't need associative relationships between child objects – the objective is clearly classifying the individual components of a building without complex parent and many to many relationships as used in thesauri.
4. Concepts for hierarchical categorization preferred to be according to general criteria and external characteristics not based on internal inherent characteristics as per taxonomies
5. Inclusion of metadata – as per ontologies hence a merge between classification and ontology schemes.

IV. CASE STUDY

An example of a current case study within this research is Toson Palace, Egypt. A full 3D point cloud scan was performed for the palace internally and externally. From this, the main existing architectural, structural and services heritage components within the palace were identified, and hence classified into 4 major ontological classes, each sub-divided into a further 4 levels of subclasses as follows:

Assembly Category →

- Orientation →

- System →
    - Type →
        - Composition

Table 1 below shows the different components aligned with the classes and subclasses (as populated from Toson Palace as an example case-study), with only an example of what is inside the 4th level of subclass "composition". The full visualisation of all classes and sub classes was created using the open source software Protégé and can be seen in Figure 1 below as per the colour codes in the table.

As can be seen in Fig 2 below, the main 4 assembly classes in the centre are those proposed to divide up a Heritage asset's components into – namely:

1. Structural components
2. Attached Architectural Components
3. Independent components
4. Cladding

This heritage classification system has opted to focus on the components only and their individual characteristics and categorisations and not allude to environmental hierarchy of space function such as in Uniclass systems (complex, entity, activity, space) due to the fact that in new builds this can be of high importance since the new buildings are in current usage as per their original design and might still need operational management related to that. However with heritage assets, the function of the building might have changed over time, or has just become an artefact, and there might not be any more connection with other buildings on the same site as per originally creating the asset due to demolition or change of environment, hence the focus is purely on the components themselves and their characteristics. These environmental aspects have been rather placed as data properties and not defining hierarchical classes.

Furthermore, the green subclasses in Fig 2 represent the orientation of the component whether vertical or horizontal. The next level of class hierarchy belongs to the system that the components fall under. This is followed by the component type in the next stage of class hierarchy (in orange) e.g. columns joists, sanitary fittings. The 5th final subclass category is the composition or material of the component to make it easier to replace.

This is part of the ontological classification system. The next step is to add

1. ***The data properties*** which describe the common attributes for instances of a class i.e. the relationship between instances and their data values. In this case of sub-class level 4 giving the opportunity to document all the different components with their different characteristics that are actually available onsite.
2. ***The object properties*** which describe the relationship between the instances of the different classes / sub classes and each other

The ***data properties*** that have been identified to be added to all instances of subclass level 4 "Composition" are:

- Code ID
- Architectural style
- Age
- Geometric ratio
- Origin
- Material name
- Allowed stresses / load bearing
- Construction method
- Condition (deterioration)
- Life expectancy
- Maintenance constraints
- Cultural Heritage value
- Reflectance
- Space function

These would aid in identifying the specific individual components within the different heritage sites and help create an international database to assist in locating substitute components, order / manufacture replacement objects or help in maintaining existing components. .

The ***object properties*** defining relationships between the data properties and sub classes are yet to be defined within this ongoing research. These are paramount, as explained previously, as they could change the relationship and relevance between different components from one heritage asset to another hence affect the way these components are maintained.

As can be seen this system satisfies the 5 characteristics of ontologies and classifications, outlined previously, that satisfy a coding system for heritage components, in the following way:

1. ***Non-exhaustive*** - The system relies on defined class hierarchy as opposed to component hierarchy hence allowing non-exhaustive addition of components based on the type of heritage asset classified.
2. ***Non-semantic –*** The differentiation between the different components relies on a coding system not on names
3. ***Simple parent-child relationships –*** with no multiple inheritance in the components as achieved through the current created 4 levels of classes and sub classes
4. ***Hierarchical categorisation is according to general criteria*** not inherent within the components themselves

Inclusion of metadata – The ontological classification relies on addition of data properties and object properties between the individual components adding more depth and richness in information to the heritage components that is required for its maintenance that is not available in current new component classifications with generic setup.

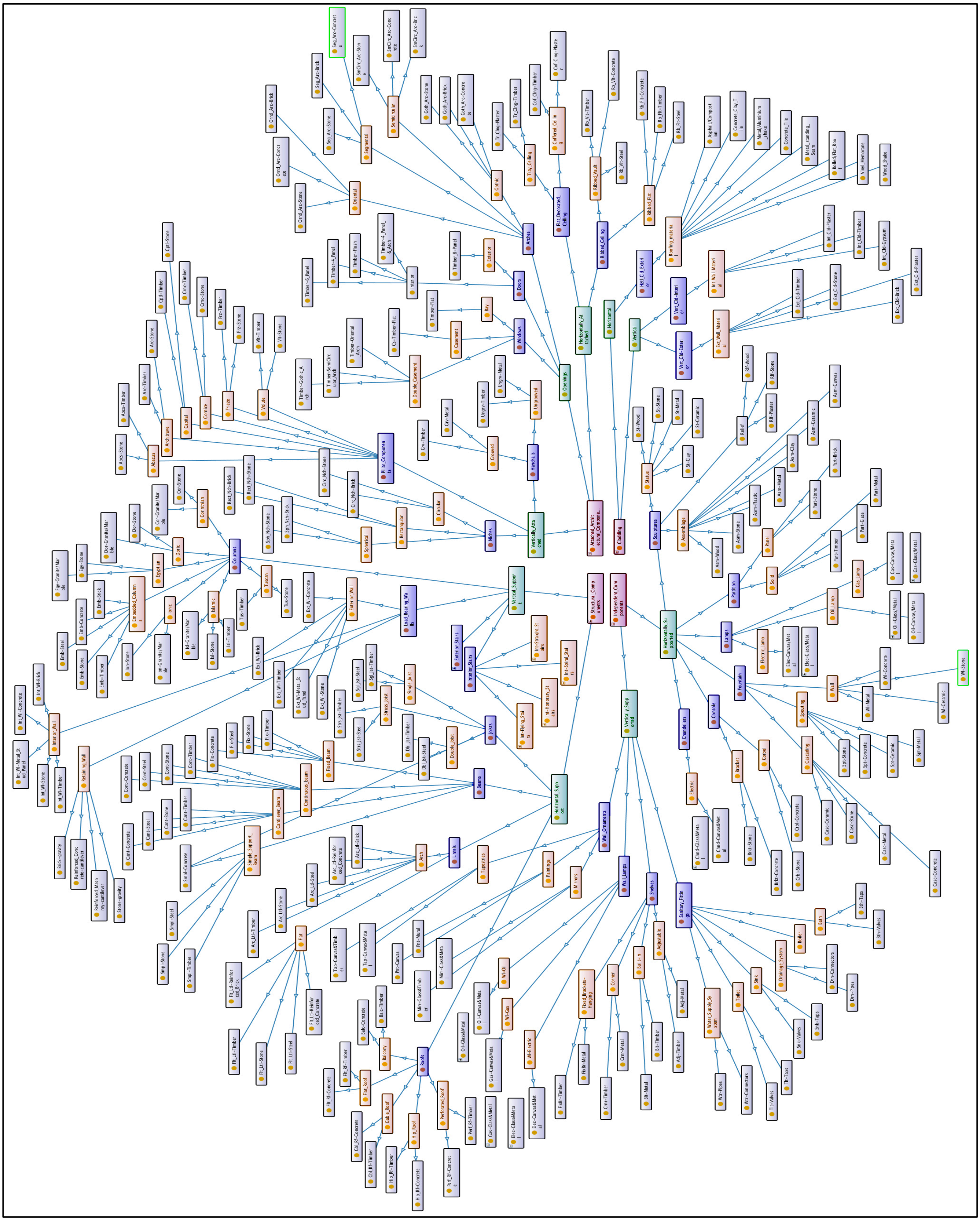

Fig. 2. Graphical Representation of the Classes and Subclasses of the Heritage Ontological Classification

TABLE 1. TABLE SHOWING CLASSES & SUBCLASSES OF THE HERITAGE ONTOLOGICAL CLASSIFICATION (EXTENDING OVER 2 PAGES)

Class 1 Assembly Category	Sub-Class2 Orientation	Sub-Class3 System	Sub-Class4 Type	Sub-Class5 Composition
Structural components	Vertical Support	Columns	Egyptian	
			Doric	
			Ionic	
			Corinthian	
			Tuscan	
			Islamic	
			Embedded Columns	Concrete
				Stone
				Brick
				Timber
				Steel
		Load bearing walls	Interior wall	
			Exterior wall	
			Retaining wall	
		Interior Stairs	Flying stairs	
			Honorary stairs	
			Spiral stairs	
		Exterior Stairs	Flying stairs	
			Honorary stairs	
			Spiral stairs	
	Horizontal Support	Beams	Simple support beam	
			Fixed beam	
			Cantilever beam	
			Continuous beam	
		Joists	Single Joists	
			Double Joists	
			Struss Joists	
		Lintels	Flat	
			Arch	
		Roofs	Flat roof	
			Gable roof	
			Hip roof	
			Shed Roof	
			Butterfly	
			Mansard	
			Gambrel	
			Perforated	
			Balcony	

Attached Architectural Components	Vertically Attached	Pillar Components	Capital	
			Architrave	
			Frieze	
			Cornice	
			Volute	
			Abacus	
		Handrails	Grooved	
			Ungrooved	
		Niches	Circular	
			Rectangular	
			Spherical	
	Horizontally Attached	Ribbed Ceiling	Ribbed Vault	
			Ribbed Flat	
		Flat Decorated Ceiling	Tray ceiling	
			Coffered ceiling	
	Openings	Arches	Semi-circular	
			Segmental	
			Gothic	
			Flat	
			Oriental	
		Doors	Interior	
			Exterior	
		Windows	Double Casement	
			Casement	
			Bay	
Independent Components	Vertically Supported	Wall Ornaments	Paintings	
			Mirrors	
			Tapestries	
		Shelves	Built in	
			Fixed Brackets - hanging	
			Corner	
			Adjustable	
		Lamps	Electric	
			Gas	
			Oil	
		Sanitary Fittings	Sink	
			Toilet	
			Bath	
			Drainage System	
			Water supply System	
			Boiler	
	Horizontally Supported	Console	Corbel	
			Bracket	
		Sculptures	Relief	
			Statue	
			Assemblage	
		Fountain	Spouting	
			Cascading	
			Wall	
		Partition	Solid	
			Panel	
		Lamps	Electric	
			Gas	
			Oil	
		Candeliers	Electric	
Cladding	Vertical	Exterior	Wall Material	
		Interior	Wall Material	
	Horizontal	Exterior	Roofing material	

REFERENCES

[1] N. Saleeb, M. Marzouk, and U. Atteya, 2018. "A Comparative Suitability Study Between Classification Systems For BIM In Heritage". International Journal of Sustainable Development and Planning, vol. 13, no. 1, pp. 130-138, 2018.

[2] A. Ekholm, "A conceptual framework for classification of construction works". ITcon Electronic Journal of Information Technology in Construction, vol. 1, pp. 1–25, 1996.

[3] K. Afsari, C.M. Eastman, "A comparison of construction classification systems used for classifying building product models". In 52nd ASC Annual International Conference Proceedings, 2016.

[4] U. Miller, "Thesaurus construction: problems and their roots". Information Processing & Management, vol. 33, no. 4, pp.481-493, 1997.

[5] J. Aitchison, & A. Gilchrist, A., "Thesaurus construction: a practical manual", (2nd ed.). London: Aslib, 1987.

[6] S. Jones, "A thesaurus data model for an intelligent retrieval system". Journal of Information Science, vol. 19, pp. 167-178, 1993.

[7] M. J. Bates, "How to use controlled vocabularies more effectively in online searching". Online, vol. 12, no. 6, pp. 45-56, 1988.

[8] R. R. Larson, "Classification clustering, probabilistic information retrieval, and the online catalog". Library Quarterly, vol. 61, pp. 133-173, 1991.

[9] D. Batty, "Thesaurus construction and maintenance: a survival kit". Database, vol. 12, pp. 13-20, 1989.

[10] J. Aitchison, "Indexing languages: classification schemes and thesauri". In L. J. Anthony (Ed.), Handbook of special librarianship and information work, 5th ed., pp. 207-261. London: Aslib., 1982.

[11] L. Lalonde, "What Is the Difference Between Classification & Taxonomy?" online, retrieved 2 August 2018, https://classroom.synonym.com/difference-between-classifi cation-taxonomy-10074596.html, 2018.

[12] S. Zhang, F. Boukamp, and J. Teizer, "Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA)". Journal of Automation in Construction, vol. 52, pp. 29-41, 2015.

[13] C. Schweizer, "What is the Difference between Taxonomy and Ontology? It is a Matter of Complexity" online, retrieved 2 August 2018, http://www.earley.com/blog/what-difference-between-taxonomy-and-ontology-it-matter-complexity, Dec 2016.